

Lebherz, Carmen; Mohr, Cordula; Henning, Matthias; Sedlmeier, Peter  
**Wie brauchbar sind Hochschul-Rankings? Eine empirische Analyse**

Teichler, Ulrich [Hrsg.]; Tippelt, Rudolf [Hrsg.]: *Hochschullandschaft im Wandel. Weinheim u.a. : Beltz 2005, S. 188-208. - (Zeitschrift für Pädagogik, Beiheft; 50)*



Quellenangabe/ Reference:

Lebherz, Carmen; Mohr, Cordula; Henning, Matthias; Sedlmeier, Peter: Wie brauchbar sind Hochschul-Rankings? Eine empirische Analyse - In: Teichler, Ulrich [Hrsg.]; Tippelt, Rudolf [Hrsg.]: *Hochschullandschaft im Wandel. Weinheim u.a. : Beltz 2005, S. 188-208* - URN: urn:nbn:de:0111-opus-73979 - DOI: 10.25656/01:7397

<https://nbn-resolving.org/urn:nbn:de:0111-opus-73979>

<https://doi.org/10.25656/01:7397>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipt.de](mailto:pedocs@dipt.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Zeitschrift für Pädagogik · 50. Beiheft

# **Hochschullandschaft im Wandel**

Herausgegeben von Ulrich Teichler und Rudolf Tippelt

Beltz Verlag · Weinheim und Basel

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, bei der die einzelnen Zahlungsmodalitäten zu erfragen sind.

© 2005 Beltz Verlag · Weinheim und Basel  
Herstellung: Klaus Kaltenberg  
Gesamtherstellung: Druckhaus »Thomas Müntzer«, Bad Langensalza  
Printed in Germany

ISSN 0514-2717  
Bestell-Nr. 41151

# Inhaltsverzeichnis

## I. Strukturelle Entwicklung und Steuerung des Hochschulwesens

*Ulrich Teichler*

Quantitative und strukturelle Entwicklungen des Hochschulwesens ..... 8

*Hans Merkens*

Zur Wettbewerbsfähigkeit des Hochschulsystems in Deutschland ..... 25

*Barbara M. Kehm/Ute Lanzendorf*

Ein neues Governance-Regime für die Hochschulen –  
mehr Markt und weniger Selbststeuerung? ..... 41

*David Phillips*

Some Concerns in Higher Education in England: A Personal Note ..... 56

## II. Wissenskulturen, Studienreform und Lehre

*Dirk Rustemeyer*

Universitäre Wissenskulturen ..... 62

*Klaus-Peter Horn*

Studienreform in der Erziehungswissenschaft –  
Zwischen Pluralisierung/Diversifizierung und Zentrierung/Vereinheitlichung ..... 76

*Ewald Terhart*

Die Lehre in den Zeiten der Modularisierung ..... 87

*Bernhard Schmidt/Rudolf Tippelt*

Besser Lehren – Neues von der Hochschuldidaktik? ..... 103

## III. Erziehungswissenschaft: Berufsübergang und Fächervergleich

*Cathleen Grunert/Ivo Züchner/Heinz-Hermann Krüger/Thomas Rauschenbach*

Der Preis des Erfolgs? Hauptfach-Pädagogen und -Pädagoginnen im Beruf ..... 116

*Thomas Rauschenbach/Rudolf Tippelt/Horst Weishaupt/Ivo Züchner*

Erziehungswissenschaft im Fächervergleich ..... 136

#### **IV. Förderung des wissenschaftlichen Nachwuchses**

*Jürgen Enders*

Wissenschaftlicher Nachwuchs in Europa ..... 158

*Julia A.M. Delius/Nicola von Hammerstein*

International Max Planck Research Schools:

Neue Wege der Graduiertenausbildung ..... 170

#### **V. Evaluation und Ranking**

*Carmen Leberherz/Cordula Mohr/Matthias Henning/Peter Sedlmeier*

Wie brauchbar sind Hochschul-Rankings? Eine empirische Analyse ..... 188

*Hildegard Schaeper*

Hochschulbildung und Schlüsselkompetenzen –

Der Beitrag der Hochschulforschung zur Evaluation der Qualifizierungs-

funktionen und -leistungen von Hochschulen ..... 209

*Carmen Lebherz/Cordula Mohr/Matthias Henning/Peter Sedlmeier*

## **Wie brauchbar sind Hochschul-Rankings? Eine empirische Analyse**

### **1. Einleitung**

Bis ins Jahr 1999 ergaben Untersuchungen an Studienanfängern, dass sich diese bei der Wahl ihrer Hochschule weniger an Tradition und Reputation der Universität für das eigene Studienfach, als vielmehr an der Nähe zum Heimatort orientieren (Bayer 1999, S. 2; Fürstenberg 2000, S. 104; Lewin u.a. 1999, S. 74f.). In der Öffentlichkeit wurde zudem die Annahme vertreten, dass die deutschen Hochschulen qualitativ vergleichbar sind. Spezialisierung und Profilbildung waren wenig fortgeschritten und wurden kaum thematisiert. Diese Haltung und der damit einhergehende fehlende Wettbewerb zwischen den Hochschulen machten einen Leistungsvergleich im Sinne von Evaluation und vergleichenden Rankings nicht erforderlich.

Heute ist die unterschiedliche Qualität der Universitäten in Lehre und Forschung Gegenstand kontroverser Diskussionen. In aktuellen Ausgaben der Wochenzeitung DIE ZEIT wird z.B. ein Forschungsranking anhand der DFG-Drittmittel publiziert, Alfred Kieser, Professor der Betriebswirtschaftslehre, diskutiert Evaluationskriterien für Hochschulen, und die Bundesministerin für Bildung und Forschung, Edelgard Bulmahn, spricht sich für ein nationales Ranking der Hochschulen aus (DIE ZEIT, 28/2003 und 30/2003). Statt der ehemals postulierten Gleichheit wird die Notwendigkeit der Differenzierung diskutiert, womit das Interesse am systematischen, wissenschaftlichen Vergleich von Hochschulen in Form von Rankings steigt.

Ein Hochschul-Ranking hat zwei Ziele: zum einen entfaltet sich der erwünschte leistungsförderliche Effekt durch Wettbewerb zwischen Hochschulen nur dann, wenn Wettbewerber und Kunden über Informationen verfügen, die Vergleiche ermöglichen. Laut Kreutzmann (2001, S. 1f.) ist es eine wesentliche Aufgabe von Rankings, die Lehr- und Forschungsprozesse sowie die Ergebnisse zutreffend zu beschreiben und damit „Markttransparenz“ zu ermöglichen. Zum anderen muss es Ziel eines Rankings sein, vielfältige Informationen so zu verdichten und zu präsentieren, dass einer definierten Zielgruppe eine Entscheidung erleichtert wird. Mögliche Zielgruppen von Hochschul-Rankings sind Abiturienten mit Studienwunsch, die sich für eine Hochschule entscheiden müssen, Studierende, die einen Hochschulwechsel planen, Professoren und Dozenten, aktuelle und potenzielle Geldgeber sowie Arbeitgeber. Daraus schließen Bayer (2001, S. 4) und Fürstenberg (2000, S. 104f.), dass Wahl und Offenlegung einer bestimmten Zielgruppe notwendige Bedingung für ein gutes Ranking sind. Aus der Zielgruppe und deren Ansprüchen an Informationen ergeben sich die geeigneten Indikatoren sowie das mathematische Modell zur Erstellung einer Rangreihe. Unterschiedliche

Zielgruppen haben unterschiedliche Ansprüche an die Art der Information; ein Ranking, das seine Zielgruppe zu heterogen fasst, verliert an Wert und kann im schlimmsten Fall keiner Gruppe als Entscheidungshilfe dienen.

Die von uns verglichenen Hochschul-Rankings von STERN (2001), STERN (2002), DER SPIEGEL (1999) und FOCUS (2002) identifizieren als Zielgruppe sowohl Studierende als auch aktuelle Abiturientenjahrgänge, die sich für ein Studium interessieren und denen mittels Ranking die Wahl einer Hochschule erleichtert werden soll. Uns ist nicht bekannt, dass die entscheidungsbildenden Kriterien zur Hochschulwahl schon einmal Thema einer publizierten Untersuchung waren. Daher lautet unsere erste Fragestellung: Welche Kriterien legen Abiturienten an die Wahl ihrer Hochschule, und inwieweit werden die Rankings diesen Ansprüchen gerecht?

Die zweite Fragestellung betrifft den Vergleich der Rankings und den Grad der Übereinstimmung ihrer Urteile. Die dritte und die vierte Fragestellung sind spezieller Natur. Sie befassen sich mit möglichen Unterschieden zwischen Universitäten der neuen und alten Bundesländer sowie mit Unterschieden hinsichtlich der Reputation zwischen Studierenden- und Professorenurteil. Diese beiden Fragen werden unabhängig von einzelnen Ranking-Ergebnissen beantwortet.

Im Folgenden werden zunächst die Daten erläutert, die wir in den vier Publikationen vorfanden, sowie unser Vorgehen bei der Datenaufbereitung. Wir gehen dann im Detail auf jede der vier Fragestellungen ein. Der Artikel schließt mit einer Gesamtdiskussion der Ergebnisse.

## 2. Datengrundlage

Aus den publizierten Hochschul-Rankings von DER SPIEGEL (1999), STERN (2001), STERN (2002) und FOCUS (2002) ergibt sich folgende Datengrundlage:

### 2.1 DER SPIEGEL (1999)

An 81 Hochschulen und 12 Fachrichtungen erfragte die Zeitschrift Urteile (Umfragen) und Fakten (statistische Daten) von 12.000 Studierenden und 1.600 Professoren. Die Durchführung oblag dem Meinungsforschungsinstitut EMNID.

Neben den Fakten *Studiendauer in Semester*, *Studierende pro Professor* und *Ab- und Zuwanderung in Prozent* wurden mittels Studierenden-Umfrage die Urteile zu *Dozentenverhalten*, *Überfüllung von Lehrveranstaltungen*, *Inhalte des Studiums* und *Ausstattung* erfasst. Die Professoren wurden außerdem nach ihren Empfehlungen von Hochschulen für ihr eigenes Fachgebiet befragt. Für jedes Fachgebiet ergab sich eine Gesamtnote sowie eine Gesamtzuordnung zu einer der Kategorien „Spitzengruppe“, „Mittelfeld“ oder „Schlusslichter“. Außerdem wurden für die Kriterien aus der Studierenden-Umfrage Mittelwerte gebildet, während die Fakten zusätzlich den Kategorien „gut“, „mittelmäßig“ oder „schlecht“ zugeordnet wurden. Beispielfhaft ergab sich für das Fach Wirt-

schaftswissenschaften an der Universität Potsdam die Gesamtnote 2,73 („gut“) aus folgenden Einzelkriterien:

Fakten (in absoluten Zahlen): *Studiendauer in Semester*: 10,6 („mittelmäßig“); *Studierende pro Professor*: 38 („gut“); *Ab- und Zuwanderung in Prozent*: -56 („schlecht“).

Urteile (Bewertungen im Range von 1 [„sehr gut“) bis 4 [„schlecht“]): *Dozentenverhalten*: 2,5; *Überfüllung von Lehrveranstaltungen*: 2,4; *Inhalte des Studiums*: 2,9; *Ausstattung*: 3,0. Es wird nicht erklärt, wie sich die Gesamtnote aus den Einzelkriterien berechnet.

## 2.2 FOCUS (2002)

Das Ranking umfasste 110 Fachhochschulen und 85 Universitäten. Die Urteile von 1.000 Dekanen und Fachbereichsleitern und etwa 3.000 Hochschullehrern flossen in das Ranking ein. Der FOCUS beauftragte die GfK Marktforschung Nürnberg zur Datenerfassung bezüglich der *Reputation* der Fachbereiche und übernahm außerdem Fakten aus Erhebungen des Statistischen Bundesamtes.

Vom Statistischen Bundesamt gingen Fakten zu den Kriterien *Wissenschaftliches Personal* (hauptberufliches Personal ohne Lehrbeauftragte), *Betreuungsrelation* (Wissenschaftler pro hundert Studierende), *Studierendenanzahl*, *Frauenanteil*, *Anteil ausländischer Studierender*, *Fachstudiodauer*, *Absolventen gesamt* und *Drittmittelquote* (Verhältnis der eingeworbenen Drittmittel eines Lehr- und Forschungsbereiches zur jeweiligen Zahl der Professoren) ein. Anhand einer Professorenbefragung durch die GfK Marktforschung wurde das Kriterium *Reputation* erhoben. Eine Auswahl aus den Angaben der Hochschulen im FOCUS-Fragebogen ging als weitere Information in das Ranking ein. Details zu Aufbau, Inhalt und Auswertung des Fragebogens wurden nicht publiziert. Die Visualisierung der Ergebnisse im FOCUS variierte stark zwischen den Kriterien. So wurden Daten des Statistischen Bundesamtes als ganze Zahlen und Prozentangaben abgedruckt, *Betreuungsrelation* und *Drittmittelquote* mit „unterdurchschnittlich“, „durchschnittlich“ oder „überdurchschnittlich“ kategorisiert, während die *Reputation* ebenfalls dreistufig, aber als Kuchendiagramm („hoch“, „mittel“, „niedrig“) dargestellt wurde. Für die Universität Potsdam ergibt sich für das Fach Wirtschaftswissenschaft folgende Information:

Fakten: *Wissenschaftliches Personal*: 40; *Betreuungsrelation*: ++; *Studierende gesamt*: 1110; *Frauenanteil in Prozent*: 43,4; *Anteil ausländischer Studierender*: +++; *Fachstudiodauer in Semester*: 11,7; *Absolventen gesamt*: 33; *Drittmittelquote*: +.

Urteile: *Reputation*: niedrig.

## 2.3 STERN (2001)

Seit 1998 wurden 9.600 Professoren und über 70.000 Studierende zu unterschiedlichen Kriterien der universitären Ausbildung befragt. Die Stichprobe umfasste 100 Hochschu-



len und 13 Fachgebiete. Durchführendes Institut für das Hochschul-Ranking des STERN war das Centrum für Hochschulentwicklung (CHE). Darüber hinaus lieferten mehr als 200 Hochschulen statistische Daten an das CHE.

Die Variable *Forschung* wurde auf unterschiedliche Weise erhoben: je nach Fachgebiet gingen die *Promotionen pro Fach*, die *Publikationen pro Professor* oder die *jährlich durchschnittlich verausgabten Drittmittel pro Lehr- und Forschungsbereich* ein. Die Variablen *Studiendauer*, *Betreuung im Praxissemester*, *Frauenanteil*, *Kontakt zu Lehrenden* und *Studienorganisation* wurden durch Befragungen von Studierenden und Professoren sowie durch statistische Daten ermittelt. Die *Ausstattung* wurde als Index aus den Fakten Arbeitsplätze, PC-Arbeitsplätze, Therapie- und Experimentalräume, Labormitarbeiter und Laborplätze sowie den Urteilen zu Audiovisuelle Medien, Bibliothek, PC-Arbeitsplätze, Räume und Labore angegeben. In welcher Aggregation diese Indikatoren in den Gesamtindex verrechnet werden, bleibt unklar, weil die Indikatoren fachspezifisch kombiniert wurden (z.B. wurden Therapie- und Experimentalräume nur für die Psychologie erhoben). Die *Ausstattung* wurde darüber hinaus nicht bei allen Fächern erhoben, sondern nur bei neun der 13 Fachrichtungen. Die Variable *Professorentipp* wurde anhand zweier Fragen erhoben: „Gesetzt den Fall, Sie haben eine Tochter oder einen Sohn, die/der Ihr Fach studieren möchte, welche drei Hochschulen in Deutschland würden Sie ihr/ihm empfehlen, wenn allein die Qualität der Ausbildung ausschlaggebend wäre?“ sowie „Welche drei Hochschulen sind nach Ihrer Einschätzung in Ihrem Fach insgesamt gesehen führend in der Forschung?“. Dabei bleibt die Verrechnung der Ergebnisse dieser Fragen zum Index *Professorentipp* ebenfalls unklar. *Betreuung* wurde sowohl als Studierenden- und Professorenurteil als auch mittels Betreuungsrelation erhoben. Der Algorithmus für den Index *Betreuung* wird nicht genannt. Die Information liegt dem Leser des Printmediums in folgender Form vor, auch hier dargestellt am Beispiel der Betriebswirtschaftslehre an der Universität Potsdam. Fakten und Urteile lassen sich hier nicht eindeutig trennen, da meist beides in den Kriterien verrechnet wurde.

*Studiendauer*: „Mittelgruppe“; *Betreuung*: „Mittelgruppe“; *Wissenschaftliche Veröffentlichungen*: „Spitzengruppe“; *Professorentipp*: „Schlussgruppe“; *Gesamturteil Studierende*: „Schlussgruppe“.

## 2.4 STERN (2002)

Auch 2002 wurde die Befragung von 10.800 Professoren und über 100.000 Studierende durch das CHE im Auftrag von STERN durchgeführt. In das Ranking aufgenommen wurden 100 Hochschulen und 22 Fachgebiete.

Für unsere Auswertungen benutzen wir ausschließlich die Print-Version. Im Vergleich zur Online-Version des Rankings bot sich eine bessere Übersichtlichkeit und einfachere Handhabung der Daten. Das Datenmaterial beider Versionen war im Hinblick auf die uns interessierenden Variablen identisch. Auf der Website des CHE ([www.che.de](http://www.che.de)) können zusätzliche Informationen abgerufen werden, die nicht in unseren Vergleich einfließen.

Für das STERN-Ranking 2002 wurden die Daten einiger Fachgebiete von Vorjahren übernommen, während andere neu erhoben wurden. Die Struktur der beiden Rankings ist identisch. Im Jahr 2002 wurde vom STERN zusätzlich eine Online-Befragung durchgeführt, in der Studierende und Professoren erneut zur Beurteilung einzelner Fächer nach vorgegebenen Kriterien aufgefordert wurden. Aus dieser Information schlossen wir, dass die Kriterien, die durch Professuren- und Studierendenurteile zustande kamen, für die Version 2002 neu erhoben wurden. Ein expliziter Hinweis dazu fand sich allerdings nicht.

### **3. Fragestellung 1: Abiturientenbefragung**

Uns interessierte, welche Kriterien Abiturienten an die Auswahl einer Hochschule legen, und ob diese Kriterien durch die vier Rankings von STERN, DER SPIEGEL und FOCUS erfasst werden. Dazu führten wir eine regionale Abiturientenumfrage durch und verglichen die Ergebnisse mit den Kriterien, die die Zeitschriften zur Erstellung ihrer Rankings benutzten. Wir erhofften uns einen ersten Hinweis darauf, wie angemessen aktuelle Rankings für die Zielgruppe der Abiturienten gestaltet sind.

#### **3.1 Methode**

Wir befragten 87 Abiturienten ( $w=45$ ,  $m=42$ ) mit Studienwunsch an einem sächsischen Gymnasium. Der von uns erstellte Fragebogen bot die Möglichkeit, einzelne Kriterien anhand einer vierstufigen Likert-Skala (1=„völlig unwichtig“, 4=„sehr wichtig“) in ihrer Wichtigkeit zu bewerten sowie weitere Kriterien zu nennen und zu bewerten. Die von uns vorgegebenen Kriterien stimmen inhaltlich mit den hierbei üblichen Kennwerten überein (Lewin u.a. 1999, S. 76f.). Wir prüften, welche der von den Abiturienten als wichtig empfundenen Kriterien (Mittelwert  $> 3.00$ , d.h. „wichtig“ oder „sehr wichtig“) auch in den Rankings thematisiert wurden.

#### **3.2 Ergebnisse**

Der Vergleich der Kriterien, die von Abiturienten an die Hochschulwahl gelegt werden und den verwendeten Kriterien in den Rankings ist in Tabelle 1 wiedergegeben.

Insgesamt sieben Kriterien erhielten eine durchschnittliche Einschätzung von mindestens 3.0. Von diesen erfasst FOCUS zwei (*Reputation*, *Betreuung*), DER SPIEGEL ebenfalls zwei (*Qualität der Lehre*, *Ausstattung*), STERN (2001) und STERN (2002) jeweils drei Kriterien (*Reputation*, *Qualität der Lehre*, *Ausstattung*). *Studentenjobs*, *Lebenshaltungskosten* und *Atmosphäre der Stadt* sind Kriterien, die ein Hochschul-Ranking schwer greifen kann. Der STERN bietet dazu in seinem Magazin „Campus&Karriere“ statistische Hintergrundinformationen zu den Universitätsstädten, die für Abiturienten

diesbezüglich aufschlussreich sein können. Als Teil eines Rankings können solche Informationen allerdings nicht unbedingt erwartet werden. Anders steht es um die *Praxisorientierung*. Diese wird von STERN lediglich für Fachhochschulen erhoben. Unsere Umfrage unter Abiturienten zeigt jedoch, dass das Interesse unabhängig davon, ob sie an Universitäten oder Fachhochschulen studieren wollen, gegeben ist (Tabelle 2).

TABELLE 1

*Abgleich der Kriterien von Abiturienten und den Rankings*

Kriterium (von den Abiturienten als „wichtig“ bzw. „sehr wichtig“ eingeschätzt)	M	Rankings, die das Kriterium erhoben		
		DER SPIEGEL	STERN (2001) und STERN (2002)	FOCUS
Ausstattung	3.48	„Ausstattung“	„Ausstattung“	
Praxisorientierung	3.44			
Qualität der Lehre	3.38	„Überfüllung von Lehrveranstaltungen“, „Studierende pro Professor“	„Kontakt zu Lehrenden“, „Studienorganisation“, „Betreuung“	„Betreuungs- relation Studierende- Professoren“
Studentenjobs	3.32			
Lebenshaltungskosten in der Stadt	3.10			
Atmosphäre der Stadt	3.06			
Ruf der Hochschule für das eigene Studienfach	3.03		„Professorentipp“	„Reputation“
<i>Anmerkungen.</i> M = Mittelwert				

TABELLE 2

*Bedeutung der Praxisorientierung für zukünftige Studierende an Universitäten und  
Fachhochschulen*

	unwichtig	wichtig	sehr wichtig	gesamt
Universität	1	29	15	45
Fachhochschule		4	7	11

### 3.3 Diskussion

Uns ist keine Studie bekannt, die in Rankings erhobene Kriterien mit den Interessen von Abiturienten abgleicht. Unsere regionale Umfrage an einem sächsischen Abiturientenjahrgang zeigt, dass insbesondere die *Praxisorientierung* in den vorliegenden Rankings vernachlässigt wird. Der STERN legt dieses Kriterium lediglich an die Fachhochschulen an; unsere Untersuchung ergab jedoch, dass auch zukünftige Studierende an Universitäten diese Information für entscheidungsrelevant halten und ihr hohe Priorität einräumen. Die Operationalisierung von Praxisorientierung ist sicherlich schwierig. Denkbar wäre zum Beispiel, Informationen zu *Studiendauer*, *Studienorganisation* und *Qualität der Betreuung im Praxissemester* der gesamten Zielgruppe aller Studierenden zugänglich zu machen. Bislang gibt es diese Informationen nur für die Fachhochschulen und ausschließlich im STERN.

Das Kriterium *Ausstattung* wird differenziert erhoben durch DER SPIEGEL und STERN, wobei der STERN zusätzlich zur Studierendenmeinung quantitative Fakten sowie die Meinung der Professoren heranzieht. Dieser multidimensionale Ansatz ist begrüßenswert, da diese Variable mehrere Aspekte umfasst und vielen Vorstellungen der Zielgruppe gerecht werden muss, z.B. zu Hardware, Räumlichkeiten, Bibliotheken oder technischen Hilfsmitteln.

Die *Betreuung* wird sehr unterschiedlich operationalisiert, jedoch durch alle Zeitschriften erhoben.

Für die *Reputation* ergab sich in unserer Umfrage, dass potenzielle Studierende differenziert nach dem Ruf ihres Studienfaches fragen, und weniger nach dem globalen Renommée der Hochschule. Diese Befunde stimmen mit den gefundenen Unterschieden zwischen Professoren- und Studierendenurteil überein, die noch dargestellt werden. Der STERN berücksichtigt dies, indem er die Meinungen zum jeweiligen Fachbereich erhebt. Der FOCUS bleibt jegliche Information schuldig, auf welcher Ebene und in welcher Form die *Reputation* erhoben wird. Da DER SPIEGEL nur Extremgruppen erhebt (besonders gute vs. besonders schlechte Universitäten), berücksichtigten wir diese Daten nicht bei unserer Analyse.

Ein Problem in Bezug auf das Kriterium *Reputation* ist, dass nicht jedes Ranking beide Gruppen, Professoren und Studierende, berücksichtigt, und dass die Ergebnisse teilweise nicht nach Fachbereichen differenziert werden.

## 4. Fragestellung 2: Übereinstimmung der Ergebnisse der Hochschul-Rankings

In unseren Vergleich gingen alle Universitäten, Musikhochschulen, Kunstakademien und privaten Hochschulen ein, die von den Zeitschriften genannt werden. Lediglich Fachhochschulen schlossen wir aus, da die Rankings an deren Bewertung teilweise andere Kriterien anlegen. Darüber hinaus war die Überlappung der befragten Fachhochschulen zwischen den Rankings gering. Die Begriffe „Hochschule“ und „Universität“ gebrauchen wir austauschbar.

Uns interessierte, inwieweit die vier Hochschul-Rankings zu gleichen Ergebnissen kommen. Dazu verglichen wir diejenigen Hochschulen, die in den Rankings jeweils die ersten zehn Plätze einnahmen, sowie beispielhaft die Ergebnisse für das Fach Betriebswirtschaftslehre (BWL). Wir haben für unsere Beispiele grundsätzlich BWL gewählt, da dieses Fach an den meisten Universitäten angeboten wird und im Vergleich zu anderen Fächern daher eine maximale Zahl von Hochschulen untersucht werden konnte. Je nach Ranking und Universität wurde BWL auch Wirtschaftswissenschaften genannt bzw. BWL und Volkswirtschaftslehre (VWL) wurden zu einem Fach BWL/VWL zusammengefasst.

#### 4.1 Gesamtvergleich der Rankings

Um die Ergebnisse der Rankings vergleichen zu können, bildeten wir für jede Universität und Zeitschrift die Mittelwerte aus allen Fächern und für alle Kriterien, die in dreistufiger Beurteilung vorlagen (z.B. „Spitzengruppe“, „Mittelgruppe“, „Schlussgruppe“).

TABELLE 3		
<i>Übereinstimmungen der Top 10 zwischen den Universitäten</i>		
Übereinstimmung	ja	nein
STERN 01 im Vergleich mit allen anderen Zeitschriften	11	19
STERN 02 im Vergleich mit allen anderen Zeitschriften	13	17
FOCUS im Vergleich mit allen anderen Zeitschriften	8	22
SPIEGEL im Vergleich mit allen anderen Zeitschriften	4	26

Der gebildeten Rangfolge entnahmen wir die Top 10 der Universitäten für das jeweilige Ranking. Wir gingen von einer Zeitschrift aus und untersuchten, ob die Universitäten aus der Top 10 dieser Zeitschrift auch unter den Top 10 der anderen Zeitschriften vertreten waren.

Es gibt keine Universität, die in allen vier Zeitschriften unter den ersten zehn Plätzen vertreten ist. Wenn man beispielsweise die Top 10 aus dem SPIEGEL mit den Top 10 aus den anderen drei Zeitschriften vergleicht, dann gibt es maximal 30 Übereinstimmungen. Tatsächlich

betrug die Zahl der Übereinstimmungen zwischen SPIEGEL und den anderen drei Zeitschriften nur vier von 30. Für die Top 10 von FOCUS fanden sich acht, für die von STERN (2001) elf und für die von STERN (2002) 13 Übereinstimmungen (Tabelle 3).

#### 4.2 Fachspezifischer Vergleich am Beispiel BWL

Für den fachspezifischen Vergleich über alle Kriterien hinweg, hier am Beispiel des Faches Betriebswirtschaftslehre (BWL) dargestellt, gingen wir gleichermaßen vor: Wir bildeten die Mittelwerte aus allen Variablen, die in dreistufiger Beurteilung vorlagen, für jede Universität und Zeitschrift, nutzten aber nur Bewertungen des Faches BWL. Auch hier bildeten wir Rangreihen aus den Mittelwerten. Der Versuch, eine Top 10 zu bilden, scheiterte daran, dass Rang 10 mehrfach durch identische Mittelwerte besetzt ist. Deshalb bildeten wir je nach Zeitschrift Rangreihen, die mehr als 10 Plätze umfassen, dabei aber als Grenzwert den Mittelwert von Rang 10 haben (Tabelle 4).

TABELLE 4				
<i>Übereinstimmungen der vorderen Plätze zwischen den Universitäten im Fach BWL</i>				
	nicht erhoben	ja	nein	gesamt
STERN (2001)				
BWL	5	12	22	39 (13 Ränge)
STERN (2002)				
BWL	3	13	25	41 (11 Ränge)
FOCUS				
BWL	10	5	20	35 (17 Ränge)
SPIEGEL				
BWL	12	13	16	41 (11 Ränge)

Im Beispielfach BWL hat DER SPIEGEL 13 von 41 möglichen Übereinstimmungen, der FOCUS fünf von 35, STERN (2001) 12 von 39 und STERN (2002) 13 von 41 Übereinstimmungen. Somit haben die Zeitschriften STERN (2001), STERN (2002) und DER SPIEGEL ungefähr ein Drittel Übereinstimmungen unter den evaluierten Rangplätzen, wohingegen der FOCUS nur für 14% seiner Hochschulen zu gleichen Ergebnissen kommt.

Da für das Fach Pädagogik/Erziehungswissenschaften die Daten aus dem Jahr 2004 im STERN vorliegen, haben wir auch dieses Fach beispielhaft dargestellt (Tabelle 5).

Es ergeben sich für die Zeitschriften STERN (2001) und STERN (2004) knappe 50% Übereinstimmungen, für den FOCUS 40% und für DER SPIEGEL 10%. Demnach kommen die verschiedenen Rankings im Fach Pädagogik zu ähnlicheren Urteilen als im Fach BWL.

TABELLE 5

Übereinstimmungen der vorderen Plätze zwischen den Universitäten im Fach Pädagogik

	nicht erhoben	ja	nein	gesamt
STERN (2001)				
Pädagogik	4	25	24	53 (13 Ränge)
STERN (2002)				
Pädagogik	3	17	32	52 (11 Ränge)
STERN (2004)				
Pädagogik	3	24	26	53 (11 Ränge)
FOCUS				
Pädagogik	7	19	21	47 (17 Ränge)
SPIEGEL				
Pädagogik	2	5	48	55 (11 Ränge)

### 4.3 Vergleich einzelner Kriterien

Wir prüften, ob die Bewertung bestimmter Kriterien in den einzelnen Studienfächern übereinstimmt. Aus den Kriterien, die von den Zeitschriften zur Evaluation herangezogen wurden, wählten wir vier für unseren Vergleich: *Ausstattung*, *Betreuung*, *Reputation* (Ruf der Hochschule für das eigene Studienfach) und *Forschung*. Von diesen waren *Ausstattung*, *Betreuung* und *Reputation* von den befragten Abiturienten als wichtige Entscheidungsgrundlage für die Hochschulwahl bewertet und von jeweils mindestens drei Zeitschriften erhoben worden. Zusätzlich untersuchten wir das Kriterium *Forschung*, sofern dieses im Sinne von Drittmitteln erhoben worden war. Da es sich bei Drittmitteln um ein Faktum handelt, das als absoluter Zahlenwert erhoben werden kann, erwarteten wir für dieses Kriterium sehr hohe Übereinstimmungen zwischen den Rankings.

Die *Ausstattung* wurde von STERN (2001), STERN (2002) und DER SPIEGEL erhoben.<sup>1</sup> Die Variable *Betreuung* fand sich in allen Zeitschriften, *Reputation* gemäß Professorenurteil in FOCUS, STERN (2001) und STERN (2002). Die Variable *Forschung* im Sinne von Drittmitteln entnahmen wir den Zeitschriften FOCUS, STERN (2001) und STERN (2002). Für jedes Studienfach und jedes Kriterium korrelierten wir die Ergeb-

1 DER SPIEGEL gab bei diesem Kriterium lediglich die Note jeder Universität an, nicht jedoch die Zuordnung zu einer der Kategorien „Spitzengruppe“, „Mittelfeld“ oder „Schlussgruppe“. Um das Kriterium dennoch entsprechend auswerten zu können, nahmen wir die Anzahl der Universitäten in Spitzen- und Schlussgruppe in der Gesamtnote als Maß und wiesen die gleiche Anzahl der besten sowie schlechtesten Noten in der Ausstattung ebenfalls diesen Kategorien zu. Alle verbleibenden Universitäten wurden dem Mittelfeld zugeordnet.

nisse der Rankings, z.B. für das Fach Pädagogik die entsprechenden Werte im Kriterium *Forschung* von FOCUS, STERN (2001) und STERN (2002).

Für die *Betreuung* ergaben sich Korrelationen zwischen  $r=.08$  und  $r=.62$ , im Durchschnitt lagen die Korrelationen bei  $r=.36$ . Den Effekt  $r=.36$  interpretieren wir als klein, da wir davon ausgehen, dass gleich lautende Kriterien das gleiche aussagen und daher nahezu identische Ergebnisse zwischen den Rankings erwarten würden. Die aufgeklärte Varianz ( $r^2$ ) beträgt hier im Durchschnitt 13%.<sup>2</sup> Abbildung 1 zeigt die Übereinstimmungen zwischen allen vier Zeitschriften in Form von Stamm- und Blatt-Diagrammen. Jede Kombination von Stamm (z.B. .6) und Blatt (z.B. 1) ergibt eine Korrelation bestehend aus zwei Dezimalstellen (z.B. .61). Jede Korrelation (Kombination von Stamm und Blatt) steht für ein Studienfach.

DER SPIEGEL & STERN (2001)	DER SPIEGEL & STERN (2002)	FOCUS & STERN (2001)	FOCUS & STERN (2002)	DER SPIEGEL & FOCUS
1.1	1.1	1.1	1.1	1.1
.9	.9	.9	.9	.9
.8	.8	.8	.8	.8
.7	.7	.7	.7	.7
1.6	.6	.6	.6	12
.5	5	.5	8	5
2.4	.4	7	.4	489
8.3	.3	99	.3	3
.2	9	71	.2	3
6.1	.1	.1	2	37
8.0	.0	8	.0	.0

Abbildung 1. Stamm- und Blatt-Diagramm zu den Korrelationen von DER SPIEGEL, FOCUS und STERN hinsichtlich *Betreuung*.

Für die *Ausstattung* lagen die berechneten Korrelationen zwischen  $r=.21$  und  $r=.59$ . Im Durchschnitt lagen die Korrelationen bei  $r=.36$ . Es handelt sich auch hier um einen kleinen Effekt. Die aufgeklärte Varianz ( $r^2$ ) beträgt im Durchschnitt 13%. Abbildung 2 zeigt die Übereinstimmungen zwischen SPIEGEL, STERN (2001) und STERN (2002).

Fächer, die bezüglich ihrer *Reputation* in die Rankings eingingen, korrelieren zwischen  $r=-.19$  und  $r=.97$ . Im Durchschnitt lagen die Korrelationen bei  $r=.71$ . Auffällig sind die Korrelationen für das Fach Pädagogik. Diese fallen mit  $r=-.14$  und  $r=-.16$  aus dem Range von  $r=.62$  bis  $r=.97$  der übrigen Fächer. Lässt man diese Extremwerte außer Acht, so steigt die durchschnittliche Korrelation auf  $r=.79$ . Insgesamt handelt es sich um einen großen Effekt, der zu einer aufgeklärten Varianz ( $r^2$ ) von 50% führt, bzw. zu 62%,

- 2 In den Sozialwissenschaften werden Effekte um  $r=.30$  in der Regel schon als substantiell betrachtet. In diesem Fall handelt es sich aber nicht um Effekte im üblichen Sinn, sondern um Reliabilitätskoeffizienten, die im Idealfall  $r=1$  sein oder sich zumindest nahe diesem Idealwert befinden sollten.



DER SPIEGEL & STERN (2001)		DER SPIEGEL & STERN (2002)
	1.0	
	.9	
	.8	
	.7	
	.6	
9	.5	48
60	.4	2
2	.3	2
9882	.2	1488
	.1	
	.0	

Abbildung 2. Stamm- und Blatt-Diagramm zu den Korrelationen von DER SPIEGEL und STERN 2001 bzw. STERN 2002 hinsichtlich *Ausstattung*

FOCUS & STERN (2001)		FOCUS & STERN (2002)
	1.0	
7	.9	7
17361	.8	25311
7335	.7	7535
2	.6	2
	.5	
	.4	
	.3	
	.2	
	.1	
	.0	
9	-.1	4

Abbildung 3. Stamm- und Blatt-Diagramm zu den Korrelationen von FOCUS und STERN 2001 bzw. STERN (2002) hinsichtlich *Reputation*.

wenn man das Fach Pädagogik außen vor lässt. Abbildung 3 zeigt die Übereinstimmung zwischen FOCUS, STERN (2001) und STERN (2002).

Für das Kriterium *Forschung (Drittmittel)* variierten die Korrelationen zwischen  $r=.33$  und  $r=.68$ . Im Durchschnitt lag die Korrelation bei  $r=.53$ . Dies entspricht einem mittleren Effekt mit einer Varianzaufklärung von durchschnittlich 28%. Abbildung 4 zeigt die Übereinstimmungen zwischen FOCUS, STERN (2001) und STERN (2002).

FOCUS & STERN (2001)		FOCUS & STERN (2002)
	1.0	
	.9	
	.8	
	.7	
85	.6	256
871	.5	168
	.4	
983	.3	99
	.2	
	.1	
	.0	

Abbildung 4. Stamm- und Blatt-Diagramm zu den Korrelationen von FOCUS und STERN 2001 bzw. STERN (2002) hinsichtlich *Forschung*.

#### 4.4 Diskussion

Der Gesamtvergleich der Rankingergebnisse hinsichtlich der Top 10-Universitäten von FOCUS, DER SPIEGEL und STERN zeigt, dass die Rankings zu sehr unterschiedlichen Ergebnissen kommen, d.h. innerhalb der jeweiligen Top 10 befinden sich unterschiedliche Universitäten. Dies ist überraschend, da die Datengrundlage hinsichtlich der evaluierten Hochschulen zwischen den Zeitschriften nahezu identisch ist. Aus Tabelle 3 ist ersichtlich, dass jeweils ein Maximum an 30 Übereinstimmungen zwischen den Zeitschriften möglich gewesen wäre, davon erreichen STERN (2001) und STERN (2002) mit elf bzw. 13 Hochschulen nicht einmal 50% Übereinstimmungen, bei DER SPIEGEL sind es nicht einmal 25%. Wir erklären uns diesen Befund mit der Unterschiedlichkeit der Kriterien auf Ebene der Fachbereiche sowie anhand unterschiedlicher Definitionen mancher Fachbereiche (z.B. Wirtschaftswissenschaften vs. BWL/VWL). Auf diesem globalen Niveau weisen die Ranking-Ergebnisse eine sehr geringe Übereinstimmung auf.

Höhere Übereinstimmungen haben wir für die fachspezifischen Vergleiche erwartet, wie hier am Beispiel der BWL behandelt. Die Übereinstimmung zwischen den Zeitschriften wurde ebenfalls anhand der Top 10-Listen erhoben. Dabei wies der STERN (2001) mit 12 von 39 möglichen Übereinstimmungen die höchste Anzahl auf, der FOCUS mit fünf von 35 die geringste. Aus methodischer Sicht sehen wir diesen Vergleich für die Zielgruppe zukünftiger Studierender maximal angemessen, da die differenzierte Betrachtung einzelner Fachbereiche aussagekräftiger ist als das Ranking ganzer Hochschulen (Bargel/El Hage 2000, S. 210). Die Übereinstimmung zwischen den Rankings für ein einzelnes Studienfach ist aber tendenziell noch schlechter als die globale Übereinstimmung. Wir nehmen an, dass sich die verwendeten Indikatoren bzw. deren gewichtete Verrechnung zwischen den Zeitschriften stark unterscheiden. Aufgrund man-

gelder Transparenz sowohl hinsichtlich der Indikatoren als auch der Aggregationen können wir dazu leider nur Vermutungen anstellen.

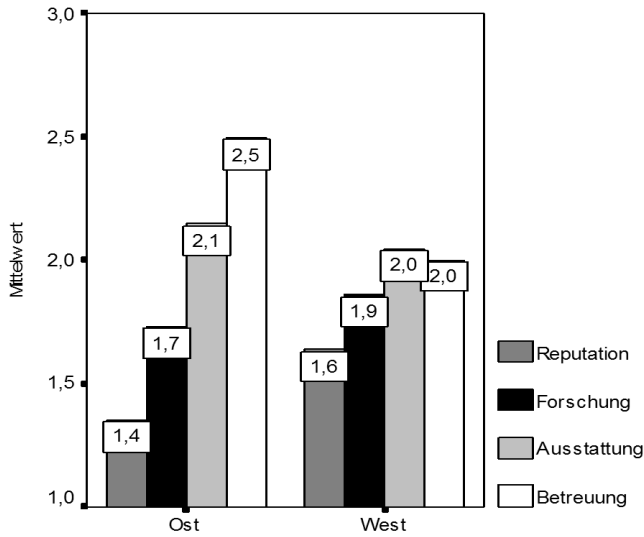
Die höchsten Ansprüche an Übereinstimmungen sind an den Vergleich einzelner Evaluationskriterien zu stellen. Hier gingen wir von der Prämisse aus, dass gleich lautende Kriterien auch das gleiche messen. In den Kriterien *Betreuung*, *Ausstattung*, *Reputation* und *Forschung* war die Wortwahl bzw. Beschreibung der Zeitschriften hinreichend identisch, um sehr hohe Übereinstimmungen naheulegen (Tabelle 1). Wir erwarteten daher für die Korrelationen der Fachbereich-Ergebnisse in jedem Kriterium Werte um  $r=.90$ . Die Effektstärken ( $r$ ) für *Betreuung*, *Ausstattung* und *Forschung* bleiben jedoch verhältnismäßig gering, sie übersteigen selten die Korrelation von  $r=.60$ .<sup>3</sup> Lediglich zwischen den Ausgaben des STERN ergeben sich hohe Korrelationen. Die Variable *Reputation*, die nur zwischen FOCUS und den Ausgaben des STERN verglichen werden können, erreichen Korrelationen ab  $r=.60$  und bis über  $r=.90$ . Beide Befunde, die geringen als auch die erwartungskonformen sind erklärbar anhand verschiedener bzw. gemeinsamer Methodenvarianz. *Reputation* wird in beiden Zeitschriften anhand einer einzelnen subjektiven Einschätzung erhoben, während die anderen Variablen Indizes aus verschiedensten, teilweise nicht benannten Einzelkriterien darstellen. Dies gilt insbesondere für die Kriterien *Betreuung* und *Ausstattung*. Vom Kriterium *Forschung* hätten wir ebenfalls hohe Korrelationen erwartet, da es sich um ein klar definiertes Kriterium handelt, die eingeworbenen *Drittmittel*.

## 5. Fragestellung 3: Vergleich der Ranking-Ergebnisse für Universitäten der alten und neuen Bundesländer

### 5.1 Methode

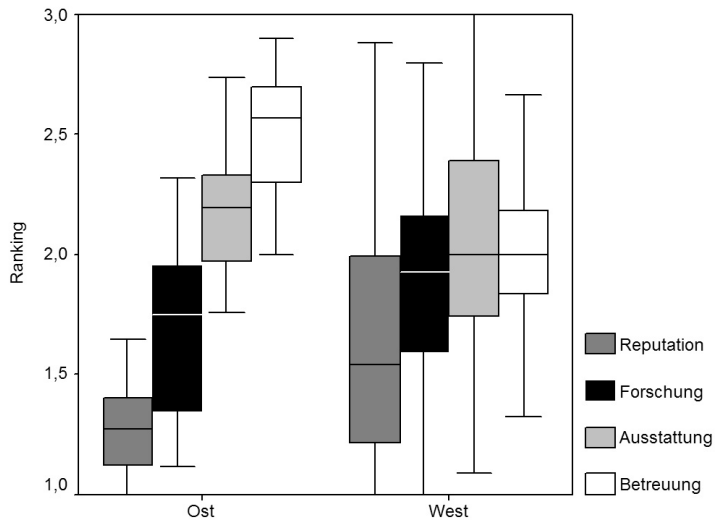
Wir teilten die Hochschulen nach ihrer geographischen Lage in die Gruppen „neue Bundesländer“ (Ost) und „alte Bundesländer“ (West) auf und verglichen die Urteile bezogen auf die Kriterien *Reputation*, *Betreuung*, *Ausstattung* und *Drittmittel* unabhängig von den einzelnen Zeitschriften. Dazu verglichen wir Mittelwerte und Mediane der einzelnen Kriterien über alle Fächer und Zeitschriften. Mittels Boxplots wurden die Verteilungen auf Extremwerte überprüft. Für alle Variablen ergaben sich 16 Universitäten der neuen Bundesländer; die *Ausstattung* war außerdem an 66 westdeutschen Universitäten erhoben worden, *Reputation* und *Betreuung* an 65 und *Forschung* (*Drittmittel*) an 60 Hochschulen der alten Bundesländer. Über die vier Zeitschriften hinweg untersuchten wir die vier Variablen in Abhängigkeit davon, ob die Universitäten sich in den alten oder in den neuen Bundesländern befinden.

3 Die Effektstärken wurden aus der Prüfstatistik  $t$  des Signifikanztests errechnet. Zur Berechnung von Effektstärken und Gründen dafür, warum sie relevanter sind als  $p$ -Werte s. z.B. Rosenthal/Rosnow/Rubin 2000, S. 9f; Sedlmeier 1996, S. 52f).



Zuordnung der Hochschulen zu neuen vs. alten Bundesländern

Abbildung 5. Vergleich der Gesamtmittelwerte der Kriterien in den neuen (Ost) vs. alten (West) Bundesländern



Zuordnung der Hochschulen zu neuen vs. alten Bundesländern

Abbildung 6. Boxplots zu den Kriterien in den neuen (Ost) vs. alten (West) Bundesländern. (Die „Boxes“ enthalten die mittleren 50% der Werte, der Strich in der Mitte gibt den Median an.)

## 5.2 Ergebnisse

Für alle vier Kriterien finden sich zumindest tendenzielle Unterschiede zwischen den Universitäten der alten und neuen Bundesländer (siehe Abbildungen 5 und 6). Der größte Effekt,  $r=.50$ , ergibt sich für das Kriterium *Betreuung* ( $t=4.08$ ,  $df=79$ ,  $p<.001$ ). Gemäß den Konventionen für üblicherweise in den Sozialwissenschaften zu erwartende Effekte (z.B. Cohen 1992, S. 157) handelt es sich um einen großen Effekt: Die Betreuung in den neuen Bundesländern wird als deutlich besser eingeschätzt als die in den alten. Der zweitgrößte Unterschied, allerdings in die entgegengesetzte Richtung, findet sich mit  $r=.26$  ( $t=1.93$ ,  $df=79$ ,  $p=.057$ ) für die *Reputation*. Demnach ist die *Reputation* an den Universitäten der alten Bundesländer höher, während Universitäten der neuen Bundesländer bessere Urteile im Kriterium *Betreuung* erhalten. Darüber hinaus sind die Effektstärken bei den Kriterien *Drittmittel* mit  $r=.22$  ( $t=-1.54$ ,  $df=74$ ,  $p=.128$ ) und *Ausstattung* mit  $r=.16$  ( $t=0.90$ ,  $df=80$ ,  $p=.371$ ) ähnlich hoch: In der *Ausstattung* schneiden die Universitäten in den neuen Bundesländern tendenziell besser ab, bezüglich *Drittmitteln* die Universitäten der alten Bundesländer.

## 5.3 Diskussion

Die Unterschiede für die Kriterien *Reputation* und *Betreuung* finden sich gleichermaßen in allen Zeitschriften. Auch bei Betrachtung der einzelnen Rankings liegen die Universitäten der neuen Bundesländer in der *Betreuung* vorn, während die *Reputation* für Universitäten der alten Bundesländer höher geschätzt wird. Als Erklärung für die bessere Betreuung an ostdeutschen Universitäten lässt sich eine verstärkte Bezuschussung in den Jahren nach der Wiedervereinigung heranziehen, die eine aktuelle Ausstattung von Bibliotheken, Hardware und Laborräumen gewährleistet, sowie die gute Betreuungsrelation durch eine relativ geringere Anzahl von Studierenden. Die *Reputation* wurde anhand des Professorentipps erhoben und weist darauf hin, dass Professoren eher Universitäten der alten Bundesländer empfehlen bzw. ihnen einen besseren Ruf zusprechen als Universitäten der neuen Bundesländer. Der Effekt der *Reputation* ist allerdings deutlich kleiner als der des Kriteriums *Betreuung*.

## 6. Fragestellung 4: Vergleich der Ergebnisse für *Reputation* zwischen Studierenden und Professoren

### 6.1 Methode

Zum Vergleich der *Reputation* im Sinne von Studierenden und Professoren verwendeten wir die Mittelwerte des *Studierendenurteils* und die Mittelwerte des *Professorentipps* aus STERN (2001) und STERN (2002). Wir wählten die Universitäten aus, die über die Zeitschriften hinweg einen Mittelwert von 2 oder höher aufweisen.

## 6.2 Ergebnisse

Die Gegenüberstellung der auf diese Art empfohlenen Hochschulen (Tabelle 6) zeigt eine sehr geringe Übereinstimmung zwischen den Urteilen von Studierenden und Professoren.

Von den 52 empfohlenen Universitäten der Studierenden finden sich nur neun unter den 18 Empfehlungen der Professoren. Der Mittelwert des *Studierendenurteils* korreliert mit dem Mittelwert des *Professorentipps* negativ mit  $r = -.13$ . Um einer möglichen Erklärung für dieses Ergebnis nachzugehen, korrelierten wir die Urteile mit dem Kriterium *Betreuung*, gemittelt aus STERN (2001) und STERN (2002). Dabei zeigte sich, dass das *Studierendenurteil* stark positiv mit *Betreuung* assoziiert ist ( $r = .61$ ,  $p < .001$ ), der *Professorentipp* dagegen negativ ( $r = -.33$ ,  $p = .004$ ).

## 6.3 Diskussion

Für das Kriterium *Reputation* konnten wir eine starke Stichprobenabhängigkeit feststellen: Die Universitäten mit „gutem Ruf“, die empfohlen werden, weisen im Vergleich von Professoren- und Studierendenurteil kaum Überlappungen auf. Auf dieses Ergebnis macht auch Bayer (1999, S. 4f) aufmerksam. Ob die Unterschiede systematisch sind, d.h. ob die Subgruppen der Empfehlungen von Professoren und Studierenden anhand bestimmter Merkmale differenzierbar sind, können wir nicht beurteilen. Korrelationen mit der Einschätzung der Betreuungsqualität sind gegeben. Eventuell ist das Kriterium *Betreuung* für Studierende der entscheidende Faktor für ausgesprochene Empfehlungen – mit den *Professorentipps* ist *Betreuung* dagegen negativ korreliert.

## 7. Gesamtdiskussion

Das Interesse an Hochschul-Rankings ist in den vergangenen Jahren stark gewachsen. Angesichts der zu erwartenden Veränderungen in der deutschen Hochschulpolitik gehen wir davon aus, dass Bedeutung und Einfluss methodisch guter Rankings in Zukunft steigen werden. Wesentliches Ziel eines Rankings muss es sein, eine bestimmte, definierte Zielgruppe mit einer angemessenen Menge an Information zu versorgen, die dieser eine Entscheidung erleichtert. Diese Aufgabe der Informationsverdichtung richtet sich jeweils an spezifische Gruppen; angesichts unterschiedlicher Interessen ist es kaum denkbar, dass ein einzelnes Ranking allen Zielgruppen, vom Abiturienten über Studierende bis zum Professor, gerecht werden kann. Dabei spielt insbesondere die Zielgruppe der Hochschulen eine Sonderrolle. Angesichts der aktuellen Diskussion um Elite-Universitäten ist es im Interesse der Hochschulen, an der Gestaltung solcher Rankings, und damit an der Wahl von Kriterien, aktiv mitzuwirken. Kriterien, die dabei vorrangig sind, unterscheiden sich jedoch maximal von denen, die für die Zielgruppe der Abiturienten oder der Studierenden interessant sind: Die Anzahl Veröffentlichungen in re-

TABELLE 6

Vergleich von Studierenden- und Professorenurteil von STERN (2001) und STERN (2002) hinsichtlich der Reputation über alle Zeitschriften hinweg

Studierendenurteil MW $\geq 2$	Professorenurteil MW $\geq 2$
Akademie der bildenden Künste Stuttgart	
Brandenburgische TU Cottbus	
E.A.P. Berlin	
European Business School Oestrich-Winkel	
Europäische Universität Frankfurt	
	FU Berlin
Hochschule der Künste Berlin	Hochschule der Künste Berlin
HH Leipzig	
	HU Berlin
Hamburger Universität für Wirtschaft	
MU Lübeck	
RWTH Aachen	RWTH Aachen
TU Bergakademie Freiberg	
TU Chemnitz	
TU Clausthal	
TU Darmstadt	
TU Hamburg-Hamburg	TU Hamburg Harburg
TU Ilmenau	
	TU München
Universität Augsburg	
Universität Bamberg	
Universität Bayreuth	
Universität Bielefeld	
	Universität Bonn
Universität Bremen	
Universität Düsseldorf	
Universität Eichstätt/Ingolstadt	
Universität Flensburg	
Universität Freiburg	Universität Freiburg
Universität Gießen	
Universität Greifswald	
Universität Halle-Wittenberg	
	Universität Heidelberg
Universität Hohenheim	
Universität Jena	
Universität Kaiserslautern	
	Universität Karlsruhe
Universität Konstanz	Universität Konstanz
Universität Landau	
Universität Lüneburg	
Universität Magdeburg	
Universität Mainz	
Universität Mannheim	
	Universität München
	Universität Münster
Universität Marburg	
Universität Oldenburg	
Universität Passau	
Universität Regensburg	
Universität Rostock	
	Universität Stuttgart
Universität Trier	
Universität Tübingen	Universität Tübingen
Universität Ulm	
Universität Vechta	
Universität Weimar	Universität Weimar
Universität Witten-Herdecke	Universität Witten-Herdecke
Uni-GH Duisburg	
Uni-GH Paderborn	
Uni-GH Siegen	
Wissenschaftliche Hochschule für Unternehmensführung Koblenz	Wissenschaftliche Hochschule für Unternehmensführung Koblenz

Anmerkungen. Hochschulen, die in beiden Gruppen einen guten Ruf genießen, sind grau hinterlegt.

nommierten Journals wie Science oder Nature etwa, die in internationalen Rankings auf die Forscher-Elite aufmerksam machen (dabei allerdings die Geistes- und Sozialwissenschaften außen vor lassen und somit naturwissenschaftlich orientierte Universitäten bevorzugen), oder das Vorhandensein von An-Instituten und die räumliche oder kooperative Nähe zu anderen Forschungsinstitutionen.<sup>4</sup>

Unsere Erhebung unter sächsischen Abiturienten hat gezeigt, dass die vorliegenden Publikationen den Kriterien der Zielgruppe der Abiturienten entgegenkommen. Lediglich die *Praxisorientierung*, ein sicherlich schwierig zu operationalisierendes Kriterium, wurde bislang vernachlässigt. Leider sprechen zwar alle Rankings Abiturienten als Lesergruppe an, schließen gleichzeitig jedoch Professoren, Studierende und weitere „stakeholders“ mit ein. Unser Vergleich von Professorenurteil und Studierendenmeinung über die zu empfehlenden Universitäten zeigt die starken Diskrepanzen auf. Demnach ist eine solche Erhebung von *Reputation* nicht geeignet, beiden Zielgruppen entscheidungsrelevante Auskunft mittels eines einzigen Rankings zu geben. Eine präzisere Definition von Personengruppen, an die sich ein bestimmtes Ranking wendet, ist bei allen vier Rankings angebracht.

Dieses Ergebnis weist auf die eingangs erläuterte Bedeutung von Zielgruppen-Definitionen hin. Angesichts der Publikationen in Wochenzeitschriften haben wir den Eindruck, dass die Zielgruppe bewusst umfassend und heterogen gehalten wurde, um einen möglichst großen Leserkreis gewinnen zu können. Darüber hinaus sehen wir keine Anzeichen, dass dieser Problematik Aufmerksamkeit gewidmet wurde, obwohl diese keineswegs neu ist (Bayer 1999, S. 4f). Unsere Ergebnisse weisen jedoch erneut darauf hin, dass ein einzelnes Ranking keine treffenden Aussagen für alle Beteiligten machen kann.

Eine ähnliche, ebenfalls methodische Kritik bezieht sich auf teilweise mangelnde Transparenz. Selbst das Ranking des STERN, das ein online abrufbares Methodenbuch zur Verfügung stellt, bietet keine Einblicke in die Algorithmen, mittels derer die Indikatoren errechnet und die Rangreihen erstellt wurden.

Unser interessantestes Ergebnis ist die überraschend geringe Übereinstimmung der Ranking-Ergebnisse. Die von uns zum Vergleich herangezogenen Rankings basieren weitgehend auf der gleichen Datengrundlage, variieren jedoch nicht nur in der Art der Darstellung für den interessierten Leser und den Erhebungs- bzw. Publikationszeitpunkten, sondern auch in ihrem Umfang an Kriterien und der Wahl und Verrechnung von Indikatoren. Wir stellten fest, dass sich die Rankings sehr uneinig darüber sind, welche Hochschulen sich die ersten zehn Rangplätze teilen, und zwar sowohl auf Ebene eines fächerübergreifenden Gesamtvergleiches, als auch bei Betrachtung einzelner Fachrichtungen, wie hier am Beispiel der BWL gezeigt. Je nach Fachrichtung variiert dabei der Grad der Übereinstimmung: so kommen die Rankings im Fach Pädagogik/Erziehungswissenschaften zu ähnlicheren Ergebnissen, mit bis zu 50% Übereinstimmung unter den Top 10-Universitäten. Selbst wenn Kriterien, die aus unserer Sicht einen ho-

4 Als Beispiel kann das Academic Ranking of World Universities der Shanghai Jiao Tong University dienen: <http://ed.sjtu.edu.cn/ranking.htm>



hen Grad an objektiver Erhebung ermöglichen (z.B. Forschung anhand von Drittmitteln), mittels identischer Fachrichtungen verglichen werden, ist die Übereinstimmung der Zuordnung zu Spitzen-, Mittel- oder Schlussgruppe gering.

Eine Ausnahme stellen die spezifischen Fragestellungen dar, die wir aufgegriffen haben: Beim Vergleich ost- und westdeutscher Hochschulen gehen die Ausprägungen der Kriterien in allen Rankings in die gleiche Richtung. Hochschulen der neuen Bundesländer bieten eine deutlich bessere *Betreuung*, während der Ruf der Hochschulen der alten Bundesländer zumindest unter den Professoren besser ist. Auch die Diskrepanzen zwischen Professoren- und Studierendenurteil finden sich Ranking-übergreifend.

Beim direkten Hochschul- und Fächervergleich jedoch kommen die Rankings nicht zu gleichen Ergebnissen. Dem interessierten Abiturienten bietet sich ein uneinheitliches Bild, das keine Entscheidungshilfe sein kann. Auch jenseits der Zielgruppendefinition zeigen sich demnach Schwierigkeiten, zu verbindlichen Aussagen zu gelangen: Für die homogene Gruppe der Abiturienten wird, bei Betrachtung mehr als eines Rankings, oftmals nur Verwirrung über unterschiedliche Empfehlungen bleiben. Wie lässt sich das erklären? Die Diskrepanzen in den Urteilen können zweierlei Ursachen haben: Die unterschiedlichen Erhebungszeitpunkte könnten dazu geführt haben, dass reale Veränderungen gemessen wurden, die sich in den divergierenden Urteilen widerspiegeln. Immerhin fanden die Erhebungen zwischen 1999 und 2002 statt. Wir vermuten jedoch, dass das unterschiedliche Vorgehen der Rankings, die uns unbekannten Aggregationen und die Art der Rangreihen-Bildung zu starker Methodenvarianz führen. Wenn diese Erklärung zutrifft, sind die Rankings in der Tat nicht vergleichbar und liefern selbst für identisch ausgeschriebene Zielgruppen sowie gleiche Fachrichtungen und Universitäten unterschiedliche Handlungsempfehlungen.

Zusammenfassend können wir zwei Schlussfolgerungen ziehen. Erstens: Ein Ranking kann nicht allen Antwort geben – wir empfehlen, Kriterien so zu wählen, dass sie für eine der Zielgruppen maximale Aussagekraft haben. Daraus folgt, dass ein Ranking, das Abiturienten die Wahl einer Universität ermöglicht, für Hochschulen und Professoren nicht gleichermaßen entscheidungsförderlich sein muss.

Zweitens: Gegen die geringen Korrelationen bei gleichen Fachrichtungen und sogar bei gleich lautenden Kriterien gibt es nur eine Lösung: Rankings müssen maximale Transparenz über Wahl und Erhebung von Kriterien gewährleisten und ihre Algorithmen zur Verrechnung dieser Kriterien offen legen.

## 8. Literatur

- Bargel, T./El Hage, N. (2000): Evaluation der Hochschullehre: Modelle, Probleme und Perspektiven. In: Zeitschrift für Pädagogik, 41, 207-224.
- Bayer, C.R. (1999): Hochschul-Ranking: Übersicht und Methodenkritik. Beiträge zur Hochschulforschung, Sonderheft 1999, Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung.
- Bayer, C.R. (2001): Eine Analyse des entscheidungsunterstützenden Informationsgehaltes deutscher und US-amerikanischer Hochschul-Rankings. In: Müller-Böling, D./Hornbostel, S./

- Berghoff, S. (Hrsg.), Hochschul-Rankings: Aussagefähigkeit, Methoden, Probleme. Gütersloh: Verlag Stiftung Bertelsmann.
- Cohen, J. (1992): A power primer. *Psychological Bulletin*, 112, S. 155-159.
- DER SPIEGEL (1999): Deutsche Hochschulen im Vergleich. Ausgabe 15/1999 (S. 58-107). Hamburg: Spiegel Verlag.
- DIE ZEIT (2003): Freiheit auf Bewährung. Interview mit Edelgard Bulmahn, Ausgabe 28/2003. Hamburg: Zeitverlag Gerd Bucerius.
- DIE ZEIT (2003): Forschung vom Fließband. Von Alfred Kieser, Ausgabe 30/2003. Hamburg: Zeitverlag Gerd Bucerius.
- FOCUS (2002): Hochschulrankings. Ausgaben 16-28/2002. Berlin: Focus Magazin Verlag GmbH.
- Fürstenberg, F. (2000): Zur Problematik von Hochschul-Rankings. In: Clemens, W./Strübing, J. (Hrsg.): *Empirische Sozialforschung und gesellschaftliche Praxis*. Opladen: Leske + Budrich.
- Lewin, K./Heublein, U./Schreiber, J./Sommer, D. (1999): *Studienanfänger im Wintersemester 1998/1999. Hochschulplanung*, Band 138. Hannover: HIS Hochschul-Informations-System GmbH.
- Kreutzmann, G. (2001): Methodik Hochschul-Ranking – Fortschritt, aber noch offene Fragen. In: *Die Betriebswirtschaft*, 61 (1), 151-152.
- Rosenthal, R./Rosnow, R. L./Rubin, D. B. (2000): *Contrasts and effect sizes in behavioral research*. New York: Cambridge University Press.
- Sedlmeier, P. (1996): Jenseits des Signifikanztest-Rituals: Ergänzungen und Alternativen. *Methods of Psychological Research - online*, 1 [Internet: <http://www.mpr-online.de/>].
- STERN (2001): Hochschulranking. *Stern spezial Campus & Karriere* (S. 38-108). Hamburg: Gruner + Jahr.
- STERN (2002): Hochschulranking. *Stern spezial Campus & Karriere* (S. 91-152). Hamburg: Gruner + Jahr.
- STERN (2004): Hochschulranking. *Stern spezial Campus & Karriere* (S. 104-166). Hamburg: Gruner + Jahr.

*Anschriften der Autorinnen und Autoren:*

- Dipl.-Psych. Carmen Lebherz, Universität Zürich, Rämistrasse 62, CH-8001 Zürich.
- Dipl.-Psych. Cordula Mohr, Drehergasse 18, 87629 Füssen.
- Dipl.-Psych. Matthias Henning, Technische Universität Chemnitz, Wilhelm-Raabe-Str. 43, 09107 Chemnitz
- Prof. Dr. Peter Sedlmeier, Technische Universität Chemnitz, Wilhelm-Raabe-Str. 43, 09107 Chemnitz